G·A·B·B·S

geospatial data analysis building blocks

Broadening Access to Geospatial Capabilities

Carol Song, Ph.D.

Senior Research Scientist

Rosen Center for Advanced Computing/ITaP

Purdue University

Purdue GIS Day
November 7, 2014

Great opening!

80/20

Open

Virtual environment

Collaborative environment

Humanities 2.0

Toolkit

Participatory

Volunteer (VGI)

Multidisciplinary

Broader community needs

Acknowledgement

- Venkatesh Merwade
- Nelson Villoria
- Lan Zhao
- Larry Biehl
- Mike McLennan
- George Howlett
- Rob Campbell
- Rajesh Kalyanam
- Many others

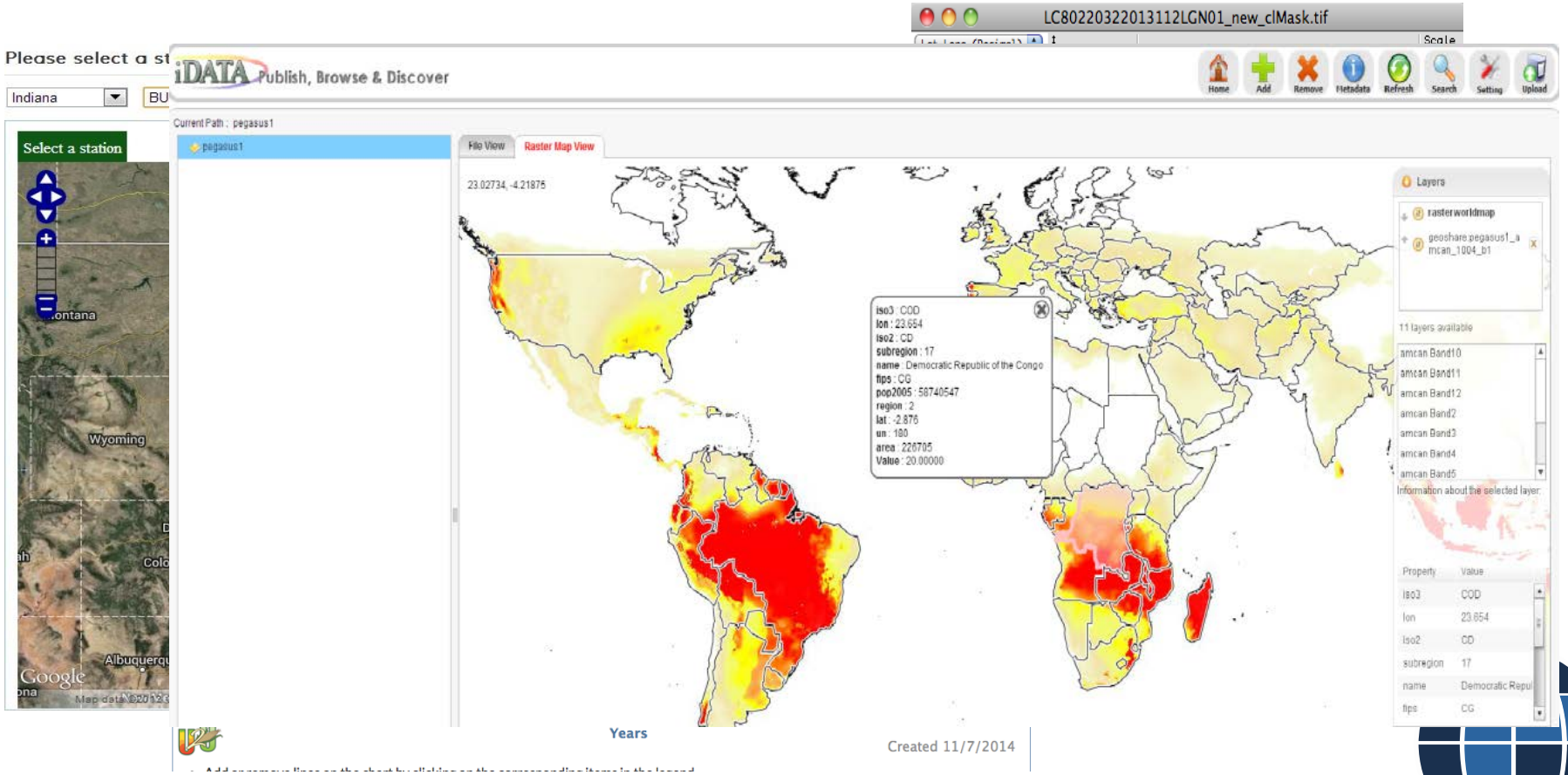
Outline

- Needs from user community
- Introduction of the GABBs project
- Technology
  - HUBzero
  - New geospatial additions
- Examples

Why do you care?

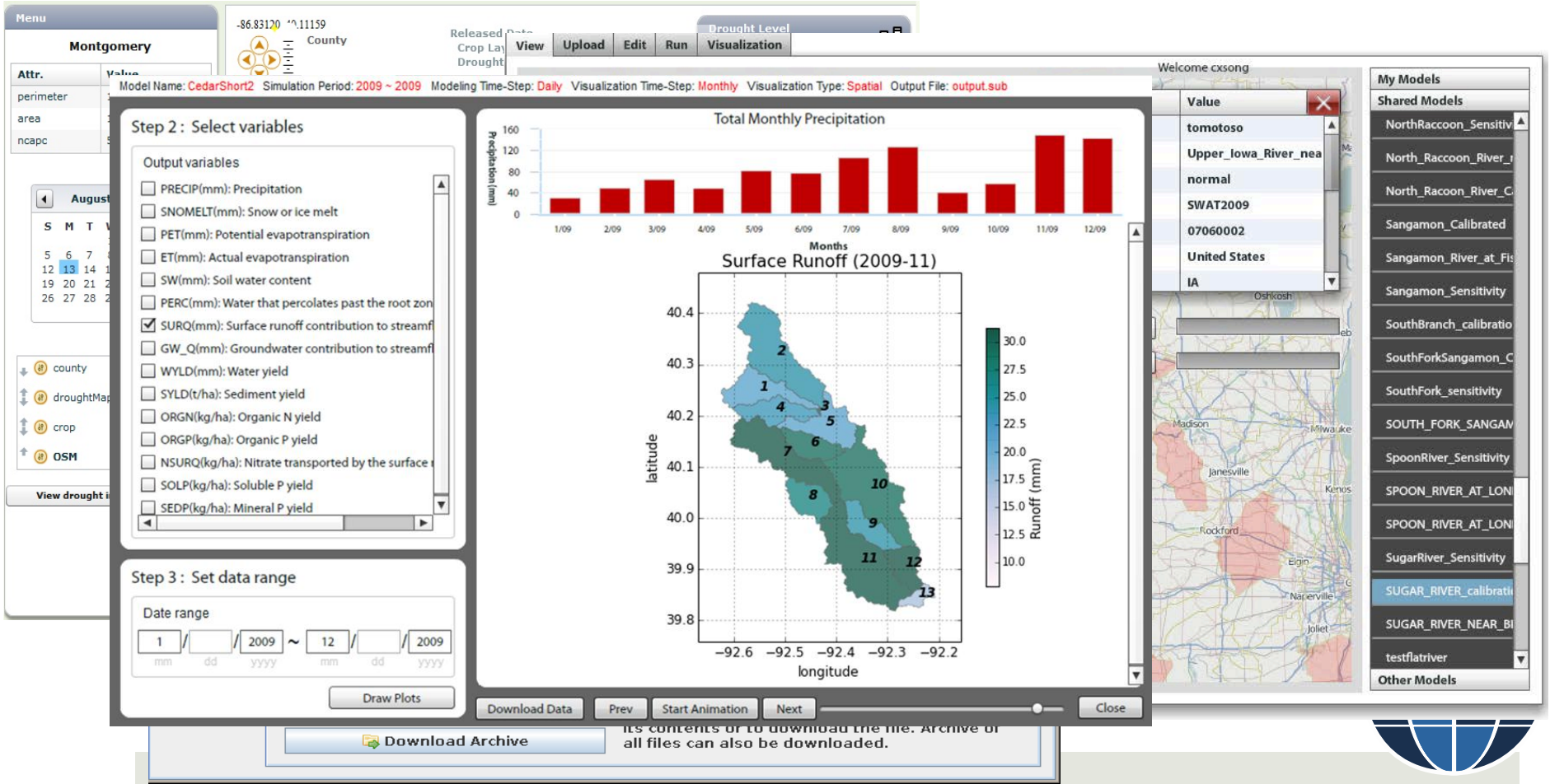
Driving Use Cases

- Easy deployment of geospatial tools



Driving Use Cases (cont'd)

- Multi-scale and multi-disciplinary data and modeling for addressing hydrologic and ag economic issues



The GABBs Project

Geospatial Modeling and Data Analysis Building Blocks

Making it easy to:

- create geospatial tools
- share these tools
- share geospatial data
- use other people's data

Key Outcomes

- The rapid tool development library RAPPTURE will support
  - geo-referenced data objects (maps, images, etc)
  - Easy way to share geospatial data, both in raw data, and visually and interactively
  - Easy way to share interactive tools that uses, and produces geospatial data
- Tool builder that supports geospatial data to further lower the barrier of creating interactive online tools
- Service interfaces to upload and share geospatial and other types of data in HUBzero
- Service interfaces to link tools and data
- Geospatial capabilities as part of core HUBzero open source release

Funding

- A National Science Foundation grant
- Data Infrastructure Building Blocks (DIBBs) program
- GABBs: 1 of 4 implementation awards in 2013
- $4.5M, 4 years (10/2013 – 9/2017)
- Collaboration with other DIBBs and DataNet awards

Team (14+)

Carol Song, PI & Project Director

Larry Biehl (CoPI, image processing and visualization)

Venkatesh Merwade (CoPI, hydrologic modeling)

Nelson Villoria (agricultural economics, sustainability)

Betsy Hillery (project manager)

Michael McLennan (HUBzero architect)

Rob Campbell (sr developer)

Leif Delgass (sr developer)

George Howlett (sr developer)

Lan Zhao (research scientist)

Rajesh Kalyanam (graduate student)

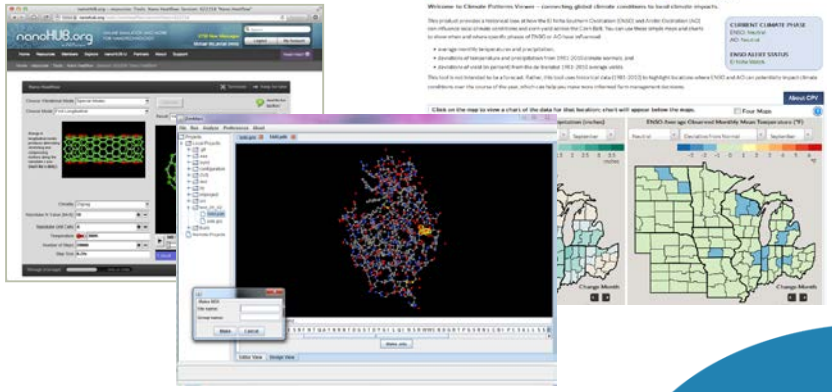
Hou-Jen Ko (graduate student)

Graduate students in scientific domains

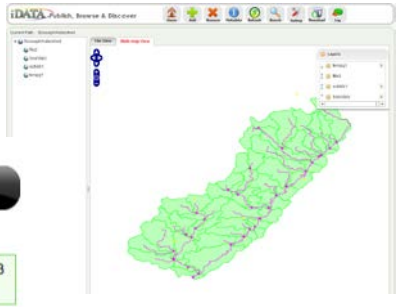
Building on prior work

- **HUBzero** (Rappture, graphics rendering, collaborative tools)
- **iData** (tool for self service data sharing and management)
- **Multispec** (tool for analyzing multispectral/hyperspectral image data)
- **Geospatial hub projects** (DRINET, Geoshare, WaterHUB, U2U etc)
- Leveraging software developed elsewhere
  - iRODS – federated data management
  - Globus – reliable transfer of large data

Platform for Scientific Collaboration

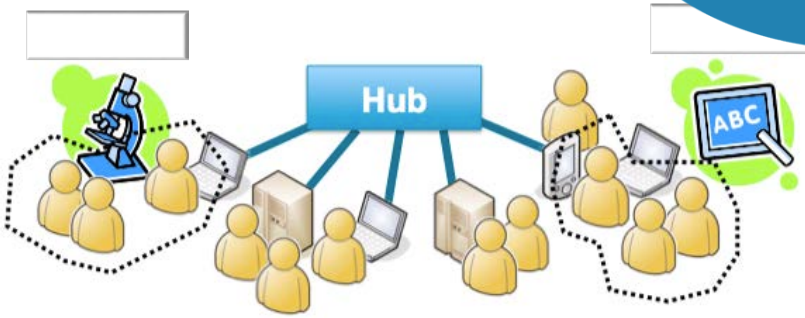


Computational Tools

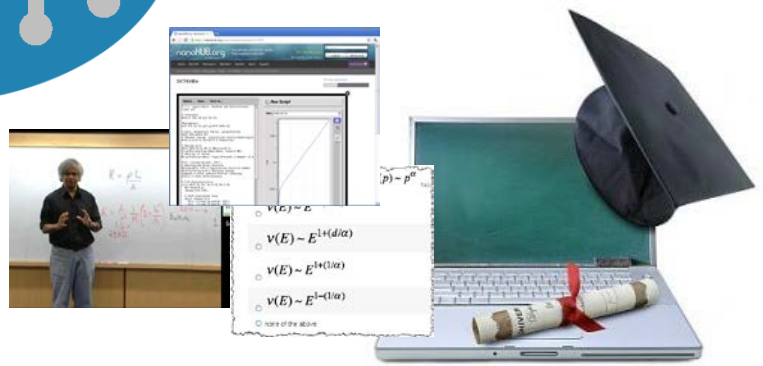


Databases / Publications

hubzero



Group/Project Collaboration



Learning Management

Courtesy of M. McLennan, Purdue University

Who's Using HUBzero?

Supporting Purdue's largest research projects:

NEES: NSF $105M - earthquake engr data (Ramirez)

NCN: NSF $18M - nanotechnology (Klimeck/Lundstrom)

C3Bio: DoE EFRC $20M - biofuels (McCann)

PRISM: DoE $17M - mems devices (Murthy/Strachan)

Supporting Purdue infrastructure

Purdue University Research Repository (PURR) – data mgmt

PurdueNExT / nanoHUB-U – online education

Supporting many other Purdue Projects

Outside Institutions

Courtesy of M. McLennan, Purdue University

60+ Hubs for many disciplines

~1,500,000
visitors total

visitors	users	
689,743	330,251	nanoHUB.org
343,350	112,862	nees.org
64,131	32,763	pharmaHUB.org
59,517	4,669	HABRIcentral.org
56,355	14,646	vhub.org
47,967	23,088	GlobalHUB.org
46,710	12,643	cceHUB.org
44,723	5,372	PURR
41,689	5,396	iemhub.org
40,289	8,207	StemEdHub.org
39,188	6,362	ciHUB.org
39,134	7,933	molecularHUB.org

Courtesy of M. McLennan, Purdue University

Tool page

AgMIP Tool: A GEOSHARE tool for aggregating outputs from the AgMIP's Global Gridded Crop Modeling Initiative (Ag-GRID)

By Nelson Benjamin Villoria, Joshua Elliott, Christoph MÃ¼ller, Jaewoo Shin[1], Lan Zhao[1]

1. Purdue University

Aggregate the yield shocks provided by the AgMIP project from their original 30x30 min resolution to any user specified level.

Launch Tool

Version **1.2.4** - published on 05 Nov 2014

Open source: license | download

View All Supporting Documents

40 users, detailed usage

Share: ...

0 review(s) (Review this)

0 questions (Ask a question)

0 wish(es) (Add a new wish)

About | Usage | Versions | Reviews | Questions | Wishlist | Supporting Docs

Category

Tools

Published on

05 Nov 2014

Abstract

As part of the Agricultural Model Intercomparison Project (AgMIP), the Global Gridded Crop Model Intercomparison group (fast track phase) have estimated historical and future changes in yields for several crops under several distinct climate change scenarios. For a detailed description of these data and its uses please refer to Rosenzweig, C. et al. (2014), Elliott, J. et al. (2014), Nelson, J. et al. (2014), and Muller and Robertson (2014).

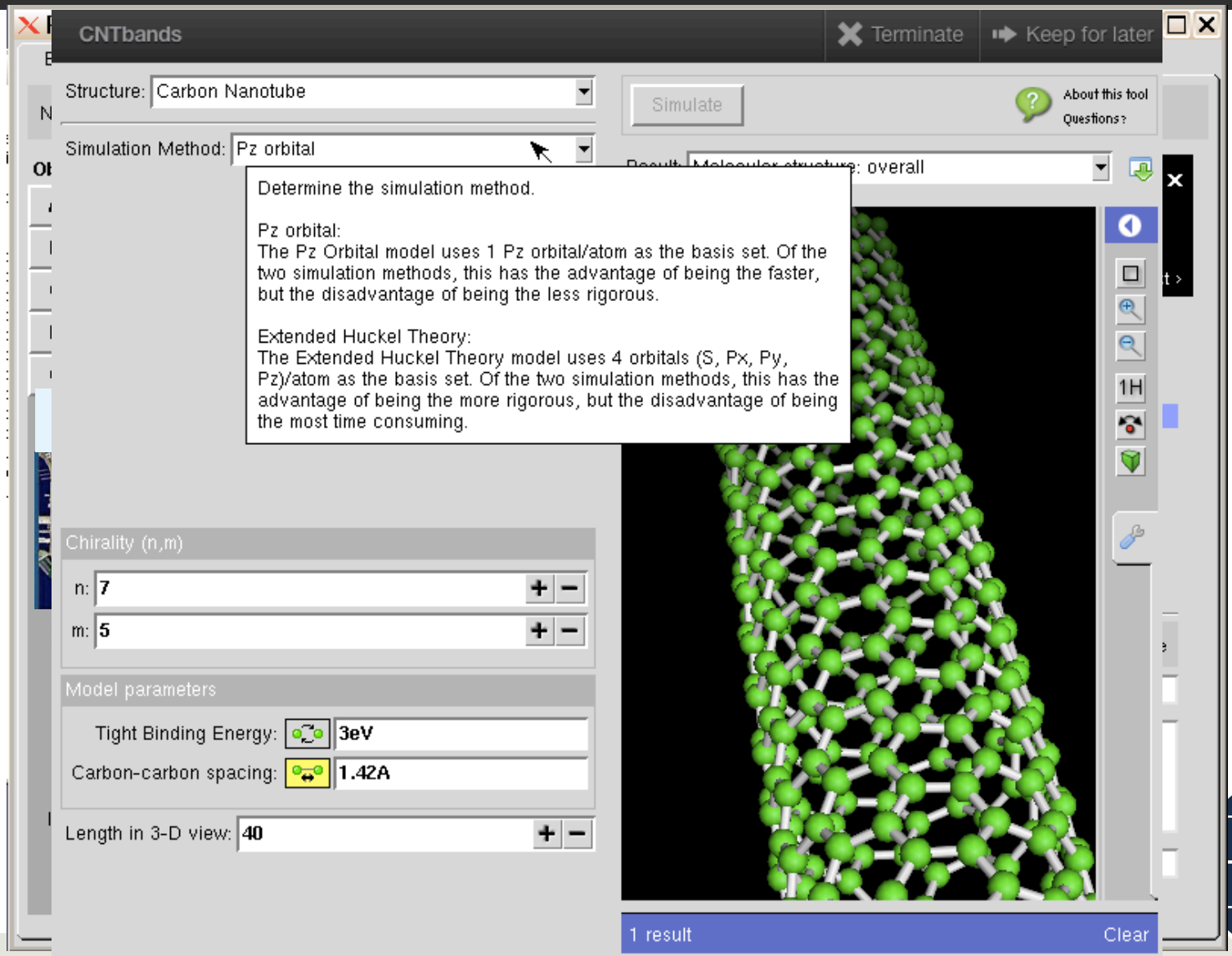
The output archive comprises time series (1971-2099) generated by seven crop models (EPIC, GEPIC, pDDSAT, LPJmL, IMAGE-AEZ, PEGASUS, LPJ-GUESS), under a number of temperature trajectories from a suite of five global climate models (HadGEM2-ES, IPSL-CMSA-LR, MIROC-EXM-CHEM, GFDL-ESM2M, NorESM1-M) and five representative

Geoshare

See also

No results found.

Rappture Builder



Publication

Localization of MLL1 morphemes in mouse mm9 genomic DNA

By Minou Bina[1], Phillip Wyss[1], Daidong Wang[1], Xiaohui Song[2]

1. Purdue University, Department of Chemistry 2. Purdue University, Academic Technologies and Rosen Center for Advanced Computing

Supplementary materials for the publication entitled "Discovery of MLL1 binding units, their localization to CpG Islands, and their potential function in mitotic chromatin." Bina M, Wyss P, et al. BMC Genomics. 14:927 (2013)

Listed in Datasets

Download (BED)

Version **1.0** - published on Mar 31, 2014
doi:10.4231/R7KW5CXF - cite this

About | Supporting Docs | Versions | Reviews | Questions | Citations

Abstract

Mixed Lineage Leukemia 1 (MLL or MLL1) plays central roles in the regulation of protein-coding genes. Previous studies of mice have established that MLL1 controls key developmental pathways including the formation of body plan during the early stages of embryogenesis. Even though MLL1 is best known for catalyzing methylation of lysine 4 in histone H3 (H3K4), reported data indicate that MLL1 performs histone methyltransferase-independent functions including gene-bookmarking during mitosis, favoring genes that were highly transcribed during interphase, to accelerate reactivation of transcription following mitotic exit (Blobel *et al*., *Mol Cell*. 2009).

By analyzing results of reported DNA binding and chromatin immunoprecipitation assays, we have uncovered the DNA sequence elements that bind the MT-domain in MLL1: Bina et. al, BMC Genomics 14:927 (2013). We describe these elements as minimal units or morphemes: the smallest 'words' in DNA that selectively bind the MLL1 MT-domain. This publication provides a file (BINA_MLL_morphemes_mm9.bed) to display the position of MLL1 morphemes in the genome browser at UCSC. You can obtain a copy of the bed file (in text format) by clicking on download. Alternatively, you can use the link below to view the position of MLL1 morphemes directly on a track in the browser at UCSC.

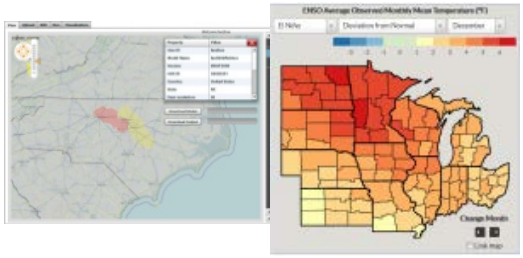
While using the genome browser, select Dense to view morpheme positions in genomic DNA. Select pack when zooming to short DNA segments. If you use the data in your research, please cite the primary publication by Bina et al., published in BMC Genomics 14:927 (2013).

Open the position of MLL1 morphemes in the mouse built mm9 genomic DNA:

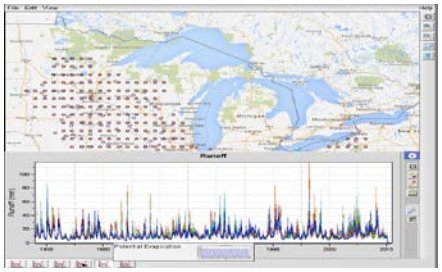
Building on HUBzero

- Extend RAPPTURE Toolkit
- Integrate map and spatial data rendering
- Adding new data objects to Rappture Builder
- Support geospatial data viewing and metadata in iData
  - Data management
  - Data sharing
  - Data quick view
  - Common processing
- Link data and tools in Hubzero
  - Dataset -> invoke tools
  - Tools -> result published and shared

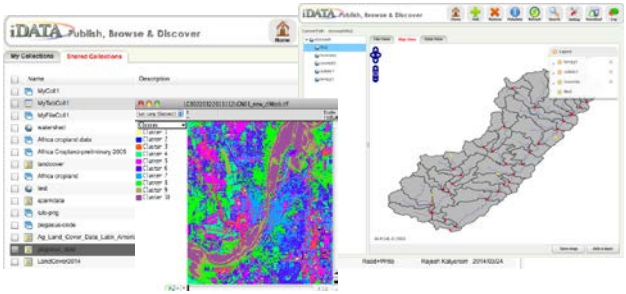
End User



Geospatial computation



Data presentation



Data Sharing

Building Blocks

Data management Rendering Tool builder Remote servers

Layers Geo-processing Data-Tool linkage

Maps Data sharing Data presentation

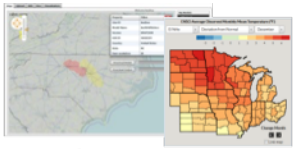
HUBzero

hubzero

Computation tools and online databases, Content publishing, Collaboration (group, project), Learning (courses, self-help), Support (tickets, Q&A), Community (forum, review, calendar)

Resources

Supercomputers, storage, map servers, geo-processing engines



Geospatial Applications

HUBzero

Web Server

Community Data Space

iDATA

Metadata Catalog

Data Manager

Non-spatial Data

Tabular Data

Raster Maps

Vector Maps

Data Services

Discover

Manage

Process

Transfer

Annotate

Visualize

WMS/WFS/WCS/ WMTS

Retrieve

SOAP/REST/ WMS/WFS/WCS/ WMTS

Publish

Visualization Server

GeoRenderer

GDAL/OGR

GEOS

OSG/osgEarth

Workspace Container

Rappture Toolkit

Image Processing API

Core API

Geospatial Mapping API

Data publishing API

Geospatial HUB Tools

iRODS

Solr

MySQL

PostGIS

Map Server

XD

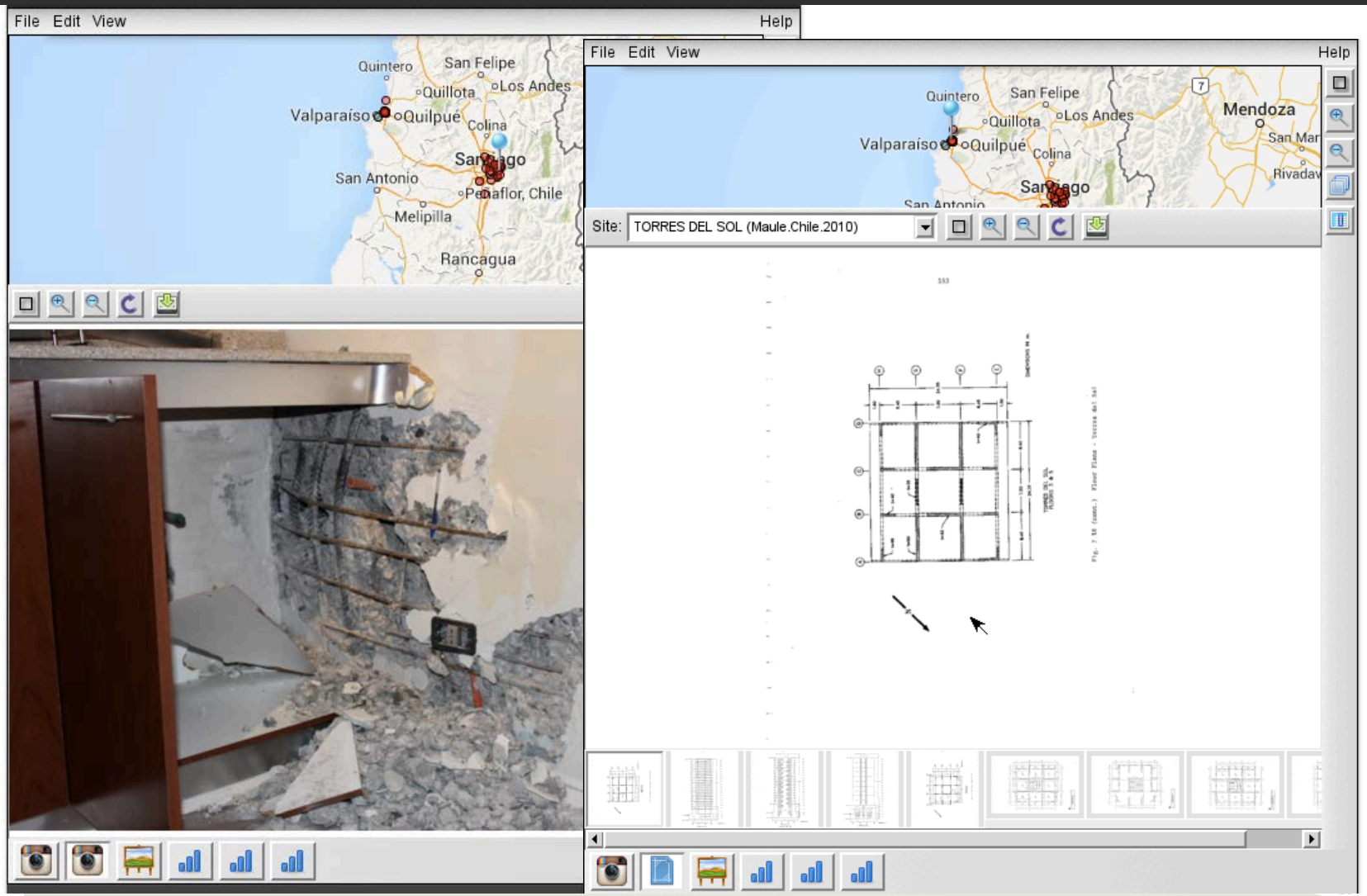
Condor

Campus Grid

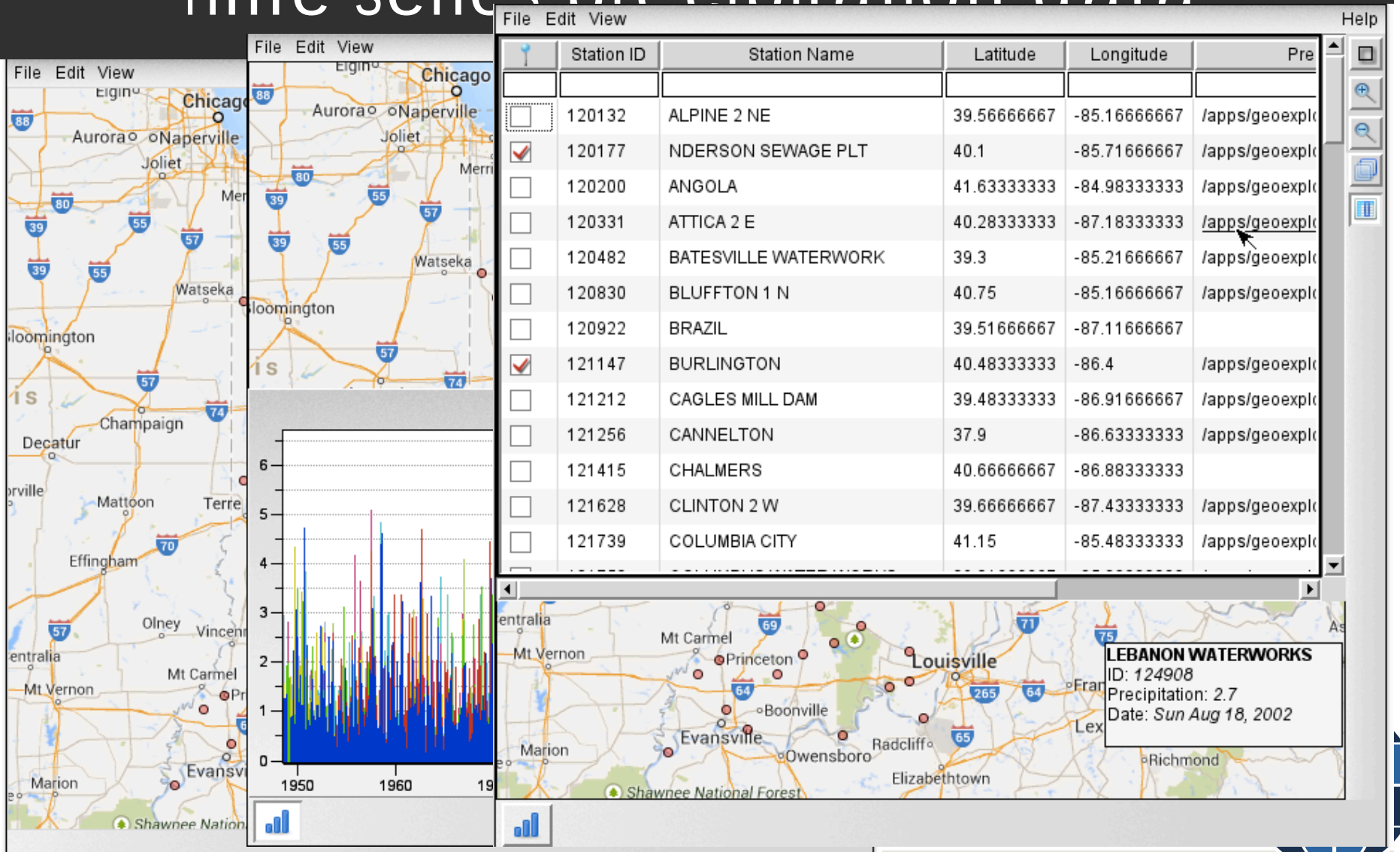
Examples

- Accessible via web browser
- Interactive
- Demonstrate capabilities

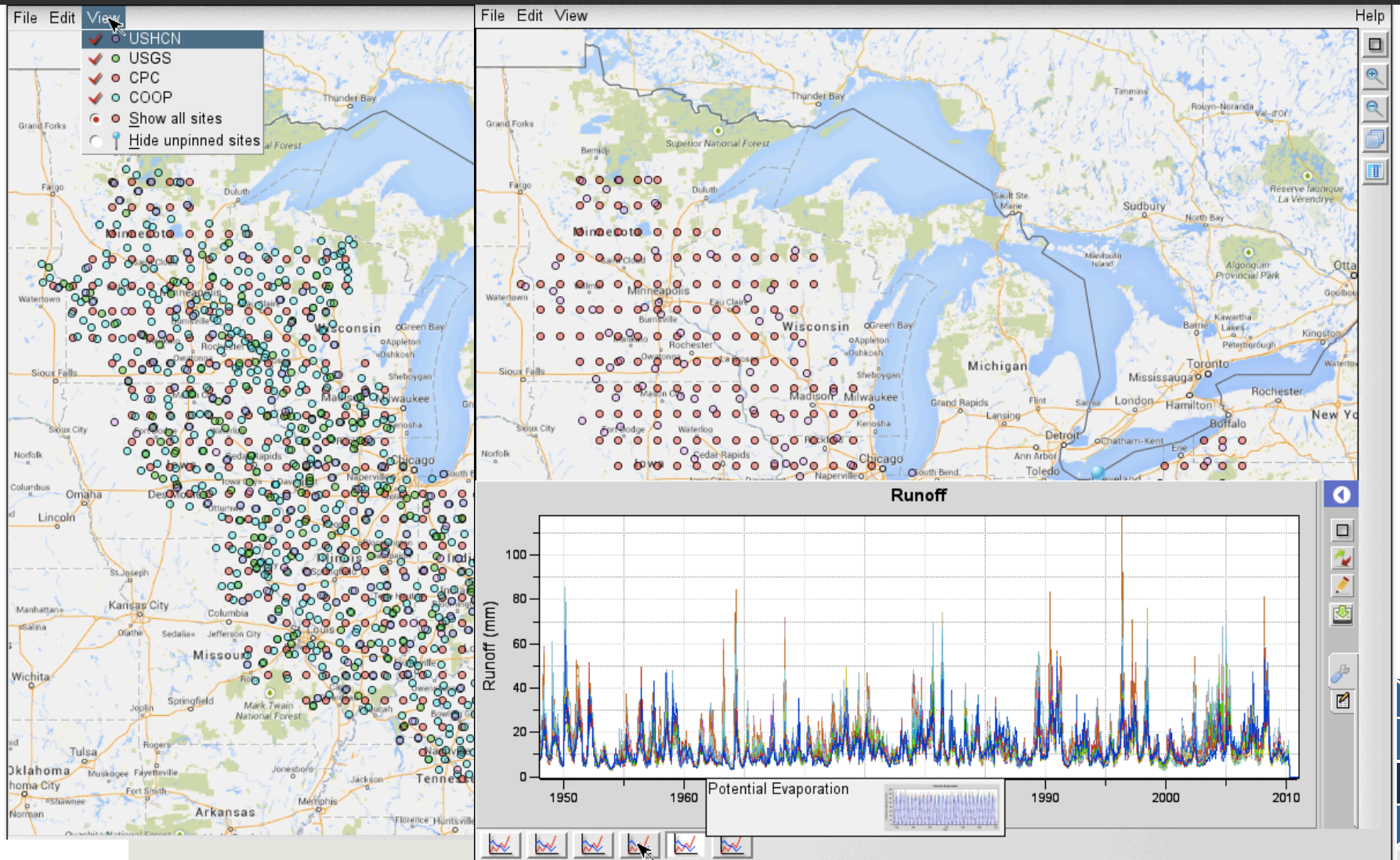
Chile Earthquake data



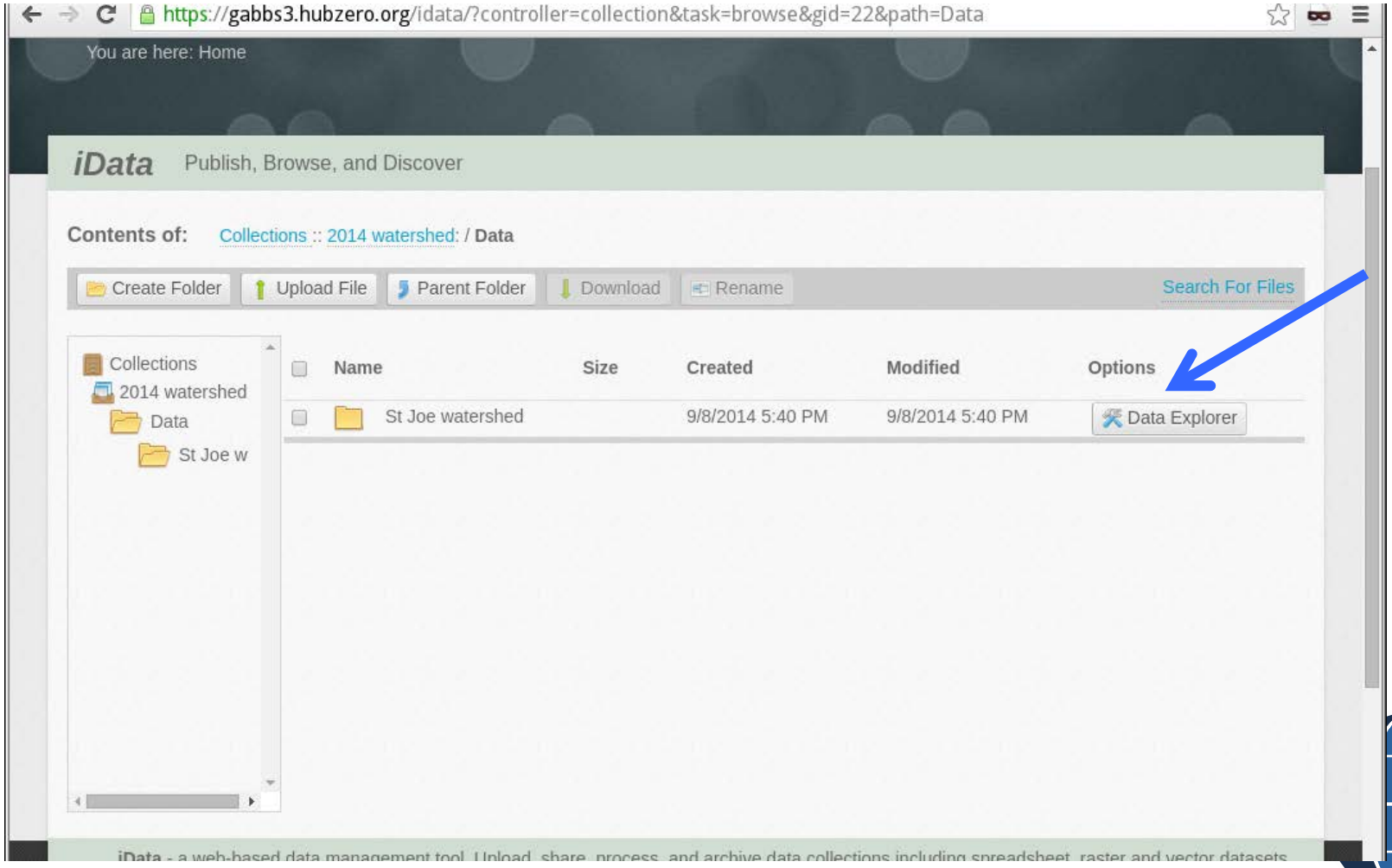
Time series precipitation data



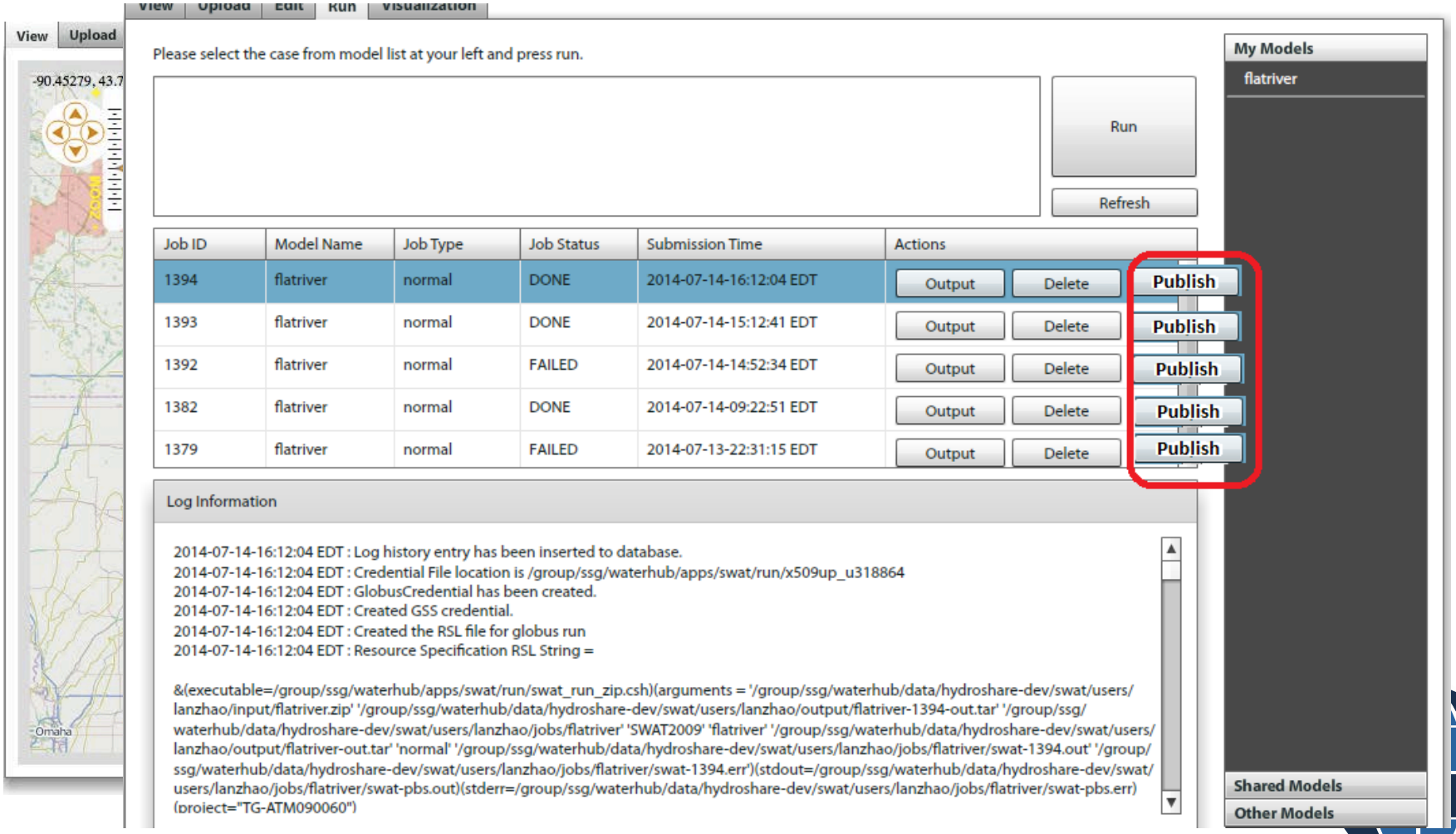
Upper MS River Basin & OH River Basin Data



Start tools from datasets

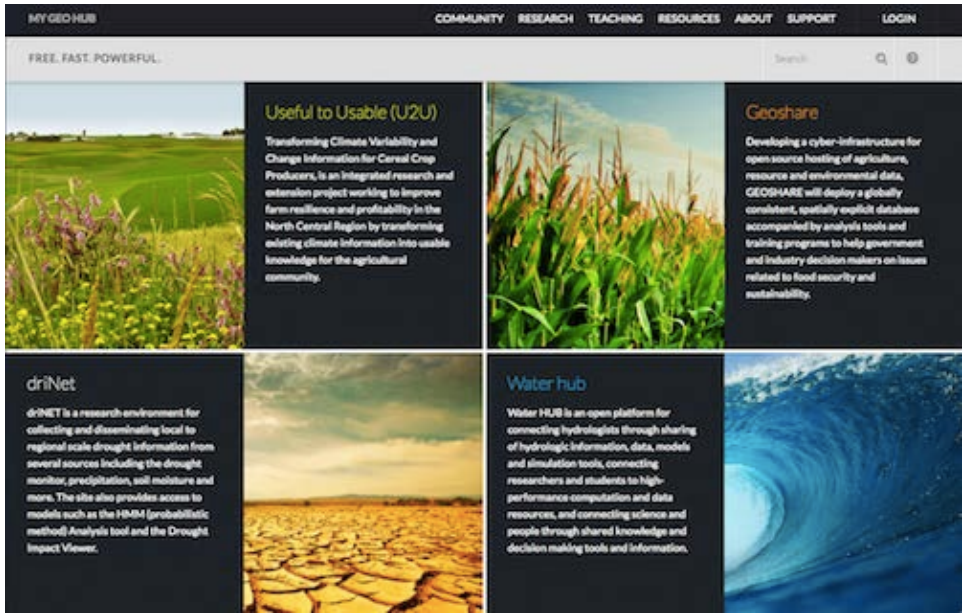


Publish data from tools

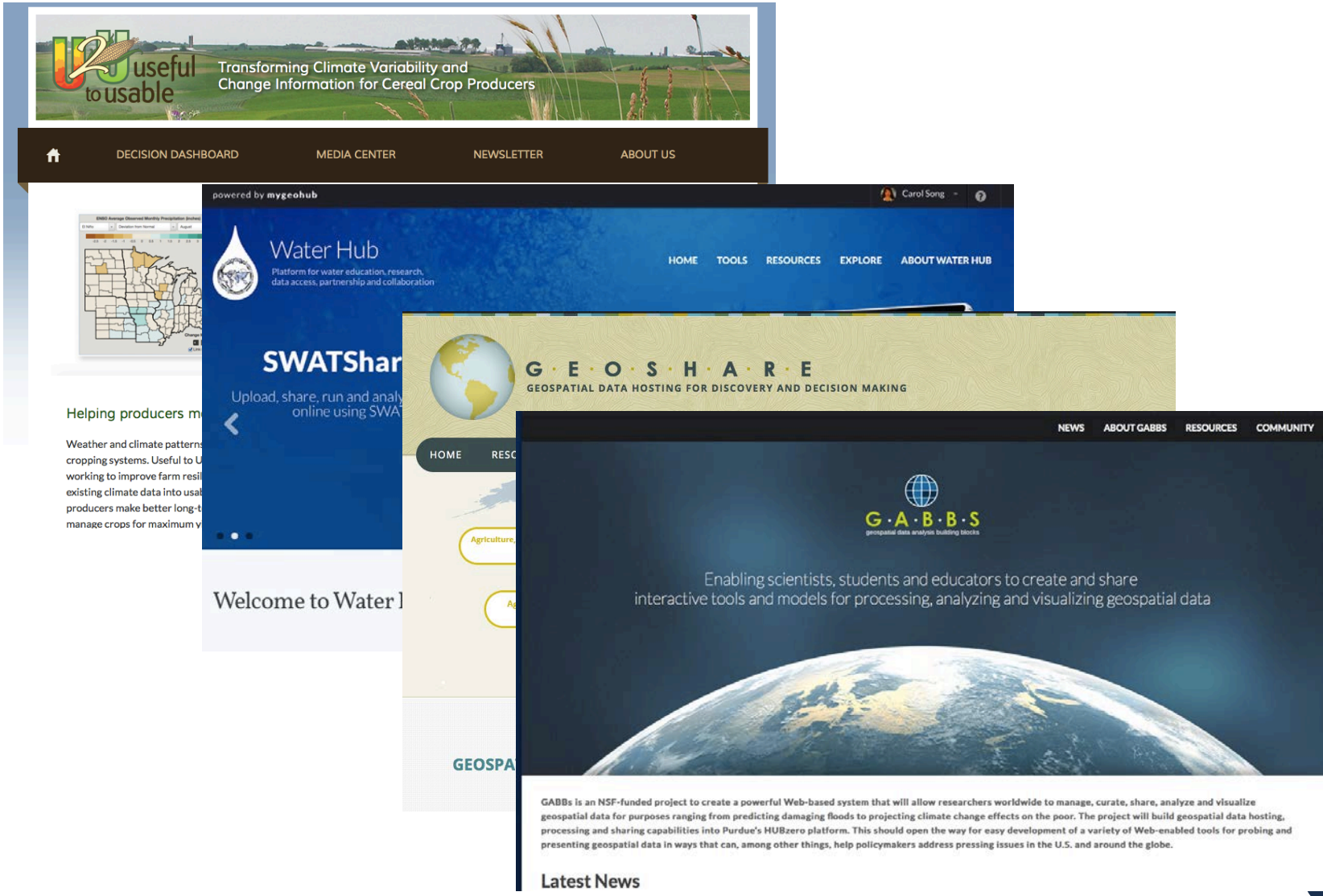


MyGeoHUB.org

- A hub for geospatial modeling, data analysis and visualization
- Hosting tools, datasets, groups and projects
- Free to sign up
- GABBs software will be installed on MyGeoHub.org
- Demonstration applications will be hosted
- User feedback and wishes are welcome!
- Build your own project, group, and community with Supergroups!



Super Groups



Collaboration

- Use cases
  - Datasets
  - applications
- Feedback, suggestions
- Participation
  - Share your work on MyGeoHub.org
- Ask questions